

ORGANIZACIÓN Y REPRESENTACIÓN DE DATOS

Introducción. - Una vez que tenemos la información representada en datos, debemos realizar un análisis estadístico con el fin de proporcionarnos utilidad a la información obtenida, en ese sentido es necesario efectuar la organización de dichos datos para posteriormente presentarlos en forma tabular y en forma gráfica. El tabular y graficar nos ayudara a poder interpretar los resultados y extraer las conclusiones pertinentes.

Puesto que tenemos dos tipos de datos: datos cualitativos (atributo), y los datos cuantitativos (variable), es necesario dar a conocer los métodos de organización para cada caso existente. Pero antes de ello veamos el formato estándar de la organización de datos.

Cuadro de distribución de frecuencias.- Los datos obtenidos mediante cualquiera de las técnicas de recolección de datos descritos anteriormente deben ser sometidos a un tratamiento y análisis estadístico, es en ese sentido que se construye un cuadro en el cual se representan las frecuencias de repetición de modalidad (categoría), del atributo o de la variable cuantitativa a consideración.

Las partes esenciales de un cuadro de distribución de frecuencias:

- a) **Numero:** Es el número de identificación del cuadro, este número se escribe a continuación la palabra cuadro
- b) **Titulo:** Es la identificación que preside al cuadro y es colocado en la parte superior del mismo:
 - Debe ser completo, es decir, debe especificar a qué se refieren los datos presentados en el cuadro, a que periodo de tiempo y el lugar al cual se refiere la información.
 - Debe ser conciso, es decir, el titulo debe ser breve y claro
- c) **Matriz de datos:** Es una matriz que contiene la información y consta de un conjunto de casillas o celdas, dispuestas en columnas y filas.

Distribución de frecuencias de un atributo

Para el caso de variables cualitativas (atributo), es decir atributos nominales y ordinales, un cuadro de distribución de frecuencias, que se construye de la siguiente forma:

Construcción de un cuadro de distribución de frecuencias

Para construir la distribución de frecuencias de un de una variable cualitativa o atributo se deben seguir las siguientes reglas o estructura:

- ✚ En la primera columna se nombra el atributo a consideración en la parte superior, a continuación las distintas modalidades existentes en dicho atributo.
- ✚ En la segunda columna se coloca la **frecuencia absoluta (n_i)**, es decir, el número de veces que se repite la modalidad respectiva a la fila en consideración del atributo.
- ✚ En la tercera columna se coloca la **frecuencias relativas (h_i)**, que puede ser expresada en forma de proporción o porcentaje

$$h_i = \frac{n_i}{n}, \quad \text{proporción}$$

$$h_i\% = \frac{n_i}{n} * 100, \quad \text{porcentaje}$$

Donde: n = cantidad de datos.

En el cuadro (3.1.), se muestra el modelo de distribución de un atributo

CUADRO (3.1.)
DISTRIBUCIÓN DE FRECUENCIAS DEL ATRIBUTO (A)
TITULO

MODALIDADES DEL ATRIBUTO A	FRECUENCIA ABSOLUTA (n_i)	FRECUENCIA RELATIVA (h_i)	FRECUENCIA RELATIVA PORCENTUAL($h_i\%$)
a_1	n_1	h_1	$h_1\%$
a_2	n_2	h_2	$h_2\%$
a_3	n_3	h_3	
.....	
a_m	n_m	h_m	
TOTAL	n	1	100

Fuente:

Remarca: Si los datos son correspondientes a **atributos ordinales**, debemos poner las modalidades en forma **ascendente o descendente**, aunque es más pragmático colocarlo en forma ascendente.

Ejemplo:. En una consultoría se ha realizado una encuesta entre 50 personas elegidas en forma aleatoria para determinar el periódico preferido de la población, obteniéndose los siguientes resultados:

P	O	T	D	T	O	R	R	P	R
R	P	R	R	P	O	D	O	T	R
O	D	T	O	T	R	T	R	R	P
R	O	P	T	D	R	D	R	P	T
R	P	T	R	O	T	O	T	D	R

En la que: T=los tiempos, R= La razón O=opinión, P= la prensa D= el deber

a) tabular u organizar los datos

Solución: a)

Cuadro # 1
Tabla de distribución de frecuencias
Periódicos de preferencia

ATRIBUTO Marcas de Periódico	n_i cantidad	h_i	h_i %
T	$n_1 = 11$	$h_1 = \frac{11}{50} = 0.22$	$h_1 \% = 22\%$
R	16	0.32	32%
O	9	0.18	18%
P	8	0.16	16%
D	6	0.12	12%
total	50	1	100%

Fuente: Elaboración propia (fuentes primarias)

¿Cómo realizamos la tabla?

En la primera columna “MARCAS DE PERIÓDICO” (se refiere al estudio al problema), y manejamos el siguiente criterio: escribimos las distintas categorías, es decir, los distintos periódicos que consideramos en el problema (solo sus abreviaturas).

Interpretación y calculo

En la segunda columna, que significa la cantidad de veces que se repite cada categoría

$n_1 = 11$, la cantidad de veces que se repite “T”

En la cuarta columna, se escribe los porcentajes de cada categoría, denotada por: $h_i \%$

$$h_i = \frac{n_i}{n} * 100, \quad h_1 = \frac{n_1}{n} * 100 = \frac{11}{50} * 100 = 22 \%$$

Realizando el mismo cálculo obtenemos:

$h_2\% = 22$, El 22% representa la categoría "T"

Ejemplo: La dirección académica de la facultad ha realizado un estudio entre los 40 bachilleres de un colegio seleccionado de forma aleatoria para determinar la carrera de su preferencia, los resultados obtenidos son:

C	I	E	C	I	E	S	C	F	C
I	F	F	C	F	F	I	C	I	S
E	S	I	F	I	C	E	S	E	F
F	C	S	S	S	I	F	F	C	S

En la que: C = Ing. Comercial, I = Informática, F = financiera, S = sistemas, E = electrónica

Construya el cuadro de distribución de frecuencias

Cuadro # 2
Tabla de distribución de frecuencias
Preferencia sobre carreras

Carreras (atributo)	n_i cantidad	h_i proporción	$h_i\%$ porcentaje
C	9	0.225	22.5%
I	8	0.2	20%
F	10	0.25	25%
S	8	0.2	20%
E	5	0.125	12.5%
total	40	1	100%

Fuente: Elaboración propia (fuentes primarias)

Remarca: Para este tipo de tabla se usará con cuatro dígitos después del punto decimal (redondeo) en la frecuencia relativa y se usará dos dígitos después del punto decimal en porcentajes (redondeo)

Interpretación

$n_3 = 10$ cantidad de estudiantes que eligen Financiera

$h_4 = 0.2$ proporción con respecto a los estudiantes de sistemas

$h_1\% = 22.5\%$ porcentaje de estudiantes que prefieren Ing. Comercial

Gráfico de representación de tablas

GRAFICO DE BARRAS.

Consiste en representar las frecuencias que corresponden a las distintas modalidades de un atributo mediante barras, rectángulos o paralelepípedos, los cuales pueden dibujarse horizontalmente o verticalmente, siendo la longitud de las barras igual a la frecuencia absoluta o relativa o porcentaje.

El grafico de barras se lo representa en un sistema de ejes coordenados, en el cual las modalidades del atributo, se representan en el eje de las abscisas y la frecuencia absoluta o relativa se representa en el eje de las ordenadas, tal como se muestran en los siguientes gráficos

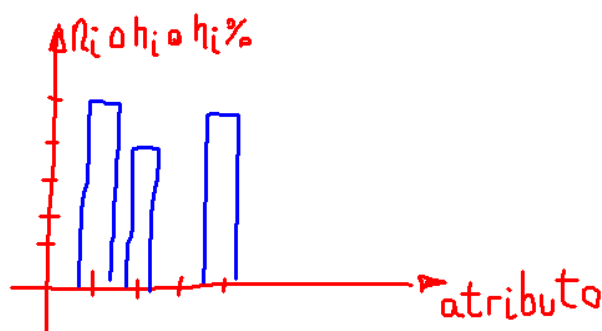


GRAFICO LINEAL

Consiste en representar las frecuencias que corresponden a las distintas modalidades de un atributo mediante puntos, los cuales pueden dibujarse horizontalmente o verticalmente, uniendo los puntos representativos con las frecuencias absolutas o relativas.

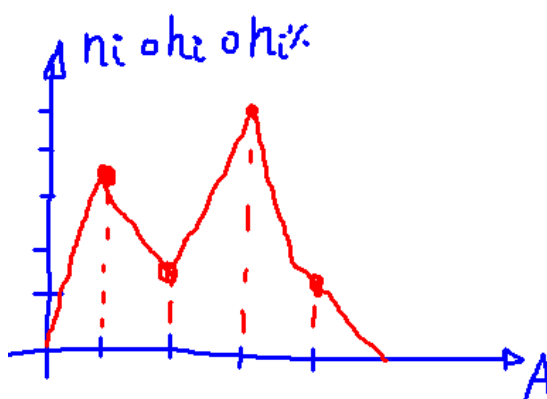


GRAFICO DE PASTEL O SECTORES (PIE)

Esta grafica es especialmente apropiada para ilustrar el espacio, cantidad o porcentaje que ocupa cada modalidad con respecto al total, con el fin de comparar modalidad y extraer peculiaridades.

Para este caso las modalidades de un atributo corresponden a cada segmento del pastel y se representan mediante la siguiente expresión matemática.

$$p_i = 360 * h_i \quad (\text{porción en grados de una modalidad respecto una circunferencia})$$

En la que:

p_i = porción del pastel .

h_i = frecuencia relativa en proporción .

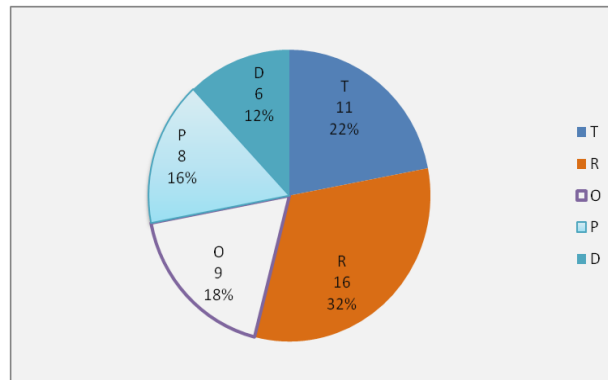
Remarca: la expresión precedente surge de realizar la regla de tres, realizado de la siguiente forma:

$$n \rightarrow 360^0$$

$$n_i \rightarrow x$$

Entendiéndose como: si el total de datos ocupa 360^0 , entonces la modalidad i, representada por la cantidad n_i , cuantos grados ocupa.

El grafico típico de pastel es:



Ejemplo:. En una consultoría se ha realizado una encuesta entre 50 personas elegidas en forma aleatoria para determinar el periódico preferido de la población, obteniéndose los siguientes resultados:

P	O	T	D	T	O	R	R	P	R
R	P	R	R	P	O	D	O	T	R
O	D	T	O	T	R	T	R	R	P
R	O	P	T	D	R	D	R	P	T
R	P	T	R	O	T	O	T	D	R

En la que: T=los tiempos, R= La razón O=opinión, P= la prensa D= el deber

a) tabular u organizar los datos

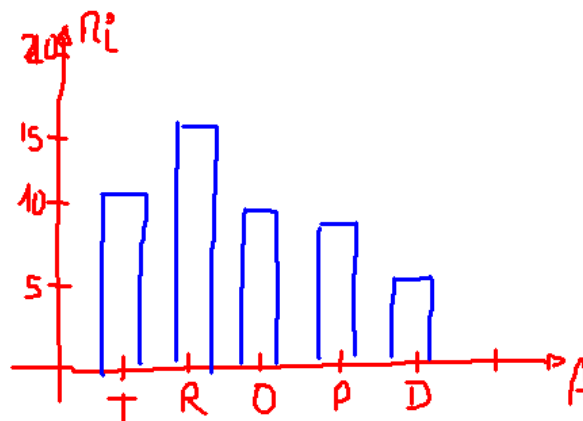
Solución: a)

Cuadro # 1
Tabla de distribución de frecuencias
Preferencia sobre periódicos en 2010

Marcas periódico (atributo)	n_i cantidad	h_i	$h_i \%$
T	11	0.22	22 %
R	16	0.32	32 %
O	9	0.18	18 %
P	8	0.16	16 %
D	6	0.12	12 %
total	50	1	100 %

Fuente: Elaboración propia (fuentes primarias)

Gráfico de barras



Realice los diagramas o gráficos

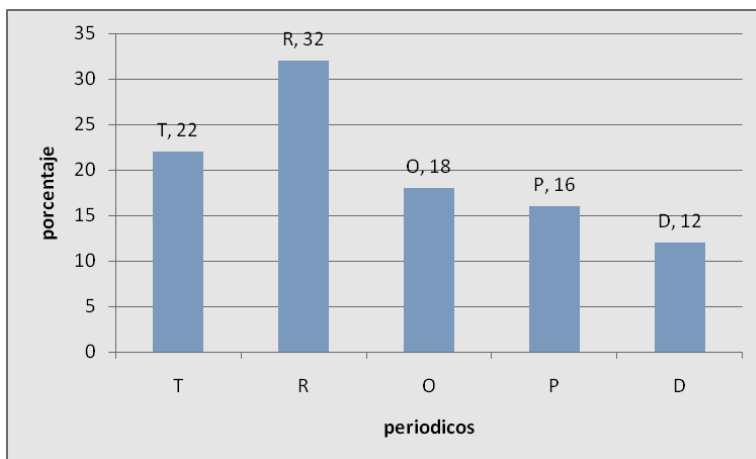


Diagrama lineal

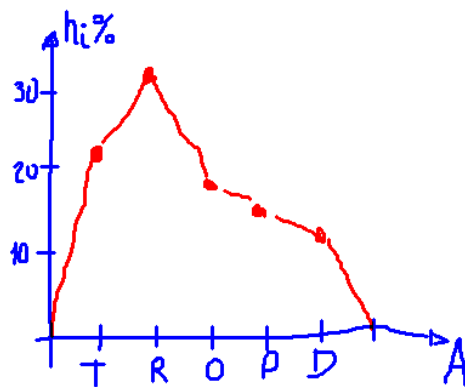


Grafico sectores

$$p_i = 360 * h_i$$

En nuestro caso tenemos:

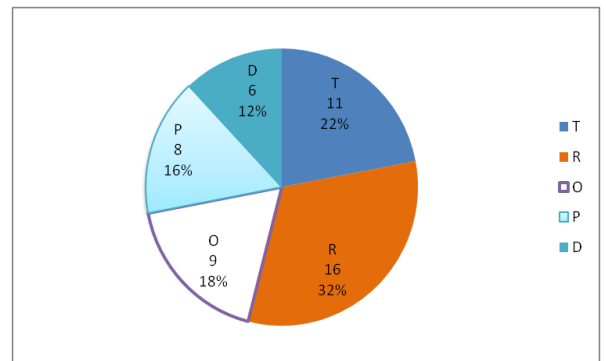
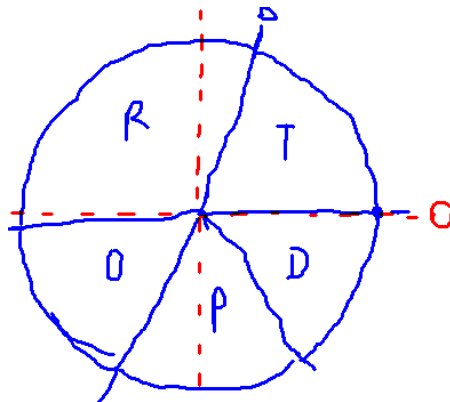
$$p_1 = 360 * h_1 = 79.2$$

$$p_2 = 115.2$$

$$p_3 = 64.8$$

$$p_4 = 57.6$$

$$p_5 = 43.2$$



Remarca: La construcción de gráficos debe ser justificada y detallada según material